

## REMARKS

This Amendment is submitted in response to the Examiner's Action dated February 25, 2003, having a shortened statutory period set to expire May 25, 2003.

In that action the Examiner has rejected claims 22-24 under 35 U.S.C. § 112, second paragraph as being indefinite for failing to particularly point out and distinctly claim the subject matter that the Applicant regards as the invention. Specifically, the Examiner notes the recitation of "the instructions" in the first line of each claim, referring back to the "computer program code" recited in the independent claim from which those claims dependent, either directly or indirectly. Based upon a careful consideration of the Examiner's comments, claims 22-24 have been amended to refer to the "computer program code" and it is believed that this Amendment overcomes the Examiner's rejection under 35 U.S.C. § 112, and withdrawal of that rejection is respectfully requested.

Next, the Examiner has rejected claims 1, 5-6, 13-15, 18-22, and 25 under 35 U.S.C. § 103(a) is being unpatentable over *Piatetsky-Shapiro*, "Discovery, Analysis, and Presentation of Strong Rules", in "Knowledge Discovery in Database", AAAI/MIT Press, 1991 (hereinafter referred to as "*Shapiro*"), in view of *Simoudis et al.*, United States Patent No. 5,692,107. That rejection is not well founded and it should be withdrawn.

The Examiner cited *Shapiro* for an alleged teaching of "comparing said one or more desired attributes and respected values with said sample population to obtain a target population." Applicants respectfully point out that the *Shapiro* reference is devoid of any teaching of the obtaining of a target population. The term "target population" is an important claim element as a statistical measure of difference between attributes and respective values

within the “target population” as compared to the “sample population” in order to reduce “the number of attributes and respective values of the sample population.”

The Examiner has stated a belief that *Shapiro* expressly teaches the obtaining of a target population at page 235 wherein *Shapiro* recites:

At the end, a cell for  $A=a$  contains the summary of all the filed tuples satisfying  $A=a$ . The summary can be presented to the user or used for deriving rules implied by  $A=a$ .

Further, the Examiner believes *Shapiro*’s teaching that the summary may preserve all of the field values and their relation to one another constitutes a teaching of the obtaining of a target population with one or more desired attributes and respective values.

Applicants urge the Examiner to consider the cited portion of *Shapiro* and the KID3 algorithm that describe the production of a summary of the sample population, not the obtaining of a target population. As noted at page 235 of *Shapiro*, *Shapiro* recites “I present here the KID3 algorithm that finds, in parallel, as simple exact rules of the form “ $(A=a) \rightarrow cond(B_i)$ ” and “. . . the cell summary is updated. . .” Applicant urges that a summary of the sample population is not a target population, having included the definitions “population”, “sample population”, and “target population” found on the Portland State University web site in a previous response. Further, as previously noted, *Shapiro* fails to teach or suggest in any way determining a statistical measure of difference between the attributes and the respective values of the target population and sample population as recited in claim 1. As set forth within the claimed invention, the selected target population is compared to the entire sample population to determine which attributes and respective values are most likely relevant in computing a predictive model. The comparison of target population to the sample population yields different results than simply reducing a dataset to a set of rules as disclosed within *Shapiro*. The results depend upon the selected target group and not the population as a whole. Different target groups

may result in different selection of the most relevant attributes. For example, a target group for a type of pizza may show a strong correlation with age and no other attribute while the target for the purchase of an expensive product may show a correlation with income.

The Examiner cites *Simoudis et al.* for its teaching of the selection of a data analysis module to perform data mining, including the use of a target population. Applicant acknowledges that *Simoudis et al.* teach the use of a target population that is employed in generating a predictive model. However, *Simoudis et al.* do not teach “comparing said one or more desired attributes and respective values with said sample population to obtain a target population” as recited in the claims of the present application. *Simoudis et al.* only teach that the target dataset typically represents a subset of an underlying data source and may be compiled from sources with data formats. (see column 4, lines 16-17) The present invention, as set forth in the claims, teaches a technique, not found in the prior art, for selecting a target group by comparing attribute values of the sample population to desired values and reducing the number of attributes by determining the statistical measure of difference between the attributes of the target and sample populations.

For an appropriate rejection under 35 U.S.C. § 103(a) the Examiner must present prior art which teaches or suggests every limitation of the claims rejected. Applicant urges that *Shapiro* and *Simoudis et al.*, whether considered alone or in combination, do not teach or suggest every claim limitation of the present invention. Specifically, this combination of prior art lacks any teaching of the determining of the statistical measure of difference between the attributes and the respective values of the target population and a sample population or comparing attributes and respective values with a sample population to obtain a target population. Accordingly, Applicant believes this rejection under 35 U.S.C. § 103(a) is not well founded and it should be withdrawn.

The Examiner has also rejected claims 3-4, 16-17, and 23-24, under 35 U.S.C. § 103(a) as being unpatentable over the combined teachings of *Shapiro*, *Simoudis et al.*, and *Dash et al.*, Dimensionality Reduction of Unsupervised Data, Proceedings, Ninth IEEE International Conference on Tools with Artificial Intelligence, November 1997. Applicant does not believe that rejection is well founded and it should be withdrawn.

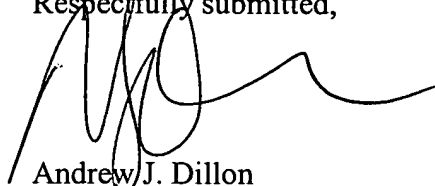
*Dash et al.*, teach a dimensionality reduction of unsupervised data utilizing an entropy measure. A sequential backwards selection algorithm, SUD, is implemented within *Dash et al.* to determine the relative importance among features by determining the relevance of particular features. *Dash et al.* is entirely silent on the subject of the reduction of variables based upon a difference between the attributes and the respective values of a target group and sample population. Consequently, Applicant urges that *Dash et al.*, whether considered alone or in combination with *Shapiro* and *Simoudis et al.* fails to teach or suggest in any way each of the claim limitations of the present application. Specifically, these combined citations lack any teachings of a determination of a statistical measure of difference between the attributes and the respective values of a target population and a sample population or the comparing of attributes and respective values with a sample population to obtain a target population. Consequently, Applicant urges that this rejection is not well founded and it should be withdrawn.

In summary, Applicant urges that, as amended herein, claims 1, 3-6 and 13-25 are clear, definite and define patentable subject over the cited prior art and withdrawal of all rejections and passage of this application to issue is respectfully requested.

No additional fee is believed to be required; however, in the event any additional fees are required, please charge IBM Corporation Deposit Account No. 09-0465. No extension of time is

believed to be necessary. However, in the event an extension of time is required, that extension of time is hereby requested. Please charge any fee associated with an extension of time to IBM Corporation Deposit Account No. 09-0465.

Respectfully submitted,

A handwritten signature in black ink, appearing to read 'Andrew J. Dillon', written over the typed name.

Andrew J. Dillon  
Registration No. 29,634  
Bracewell & Patterson, LLP  
P.O. Box 969  
Austin, Texas 78767-0969  
(512) 472-7800  
ATTORNEYS FOR APPLICANT